HEARING NANO-STRUCTURES: A CASE STUDY IN TIMBRAL SONIFICATION

Margaret Schedel

Department of Music,
Stony Brook University,
Stony Brook, NY, USA
mschedel@notes.cc.sunysb.edu

Kevin G. Yager

Center for Functional Nanomaterials,
Brookhaven National Laboratory,
Upton, NY, USA
kyager@bnl.gov

ABSTRACT

We explore the sonification of x-ray scattering data, which are two-dimensional arrays of intensity whose meaning is obscure and non-intuitive. Direct mapping of the experimental data into sound is found to produce timbral sonifications that, while sacrificing conventional aesthetic appeal, provide a rich auditory landscape for exploration. We discuss the optimization of sonification variables, and speculate on potential real-world applications.

1. INTRODUCTION

Sonification of datasets is becoming more popular as an alternative modality for exploring, and understanding, datasets. Beyond the obvious implications for accessibility, sonification enables interested parties to interact with data more deeply; e.g. multi-modal data exploration leverages more of a person's sensory 'surface area.' This is especially relevant in light of the modern trends in data collection: datasets are growing ever-larger, and in many cases ever-more complex, esoteric, and non-intuitive. We elected to study sonification of x-ray scattering data, which are rather abstract datasets that even experts struggle to understand.

An x-ray scattering experiment consists of directing a highly collimated, monochromatic, beam of x-rays through a sample of interest. The incident x-ray wave scatters off of all the atoms and/or particles in the sample, and the interference of these secondary waves produces scattered rays at angle that are characteristic of the material's internal structure. [1] In a scattering experiment, the deflection of scattered rays is characterized by the so-called *momentum transfer vector*, usually denoted by q, which is computed from the measured scattering angle, 2θ, by:

$$q = \frac{4\pi}{\lambda}\sin\theta \qquad (1)$$

where λ is the wavelength of the x-rays. The quantity q has units of 1/distance, and q-space is thus frequently called 'inverse space,' or 'reciprocal space.' This abstract space is in some sense the Fourier transform of the realspace density distribution in the sample. Mathematically:

$$s(\mathbf{q}) = \sum_n f_n e^{i\mathbf{q}\cdot\mathbf{r}} \qquad (2)$$

$$f_n(\mathbf{q}) = \int \rho(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \mathrm{d}V \qquad (3)$$

The scattered intensity, $s(\mathbf{q})$, is computed by summing the scattering contributions from the n scattering entities in the material (e.g. each atom). The scattering contribution of each entity, f_n, is in turn computed by integrating its density distribution, $\rho(\mathbf{r})$, over all of real-space.

Conceptually, the scattering experiment encodes all the information about the sample's shape and internal structure, albeit in an opaque and non-intuitive way. Roughly, a scattering peak at a particular q (i.e. angle) implies a real-space repeating structure with a size-scale of:

$$d = \frac{2\pi}{q} \qquad (4)$$

We note that the inverse nature of $2\pi/d$ means that a scattering peak at large angle corresponds to small real-space distances, whereas a peak at small angle corresponds to larger real-space distances. As the field of nanotechnology matures, x-ray scattering is emerging as a powerful tool to study new materials; however interpreting this data is difficult. Although scattering data is in essence a Fourier transform of the material's structure, an experiment only captures the amplitude of the scattered waves, and cannot record the phase information.

X-ray scattering datasets are normally visualized using two-dimensional false-color images (see Figure 1). These images are an extremely valuable tool for researchers, but have their limitations. Scattering data can have a very large dynamic range, which is difficult to represent in a single image. Here, sonification can help, since the human ear has a correspondingly large dynamic range. [2] Moreover, the Fourier transform nature of scattering data implies a natural match with audio data. In scattering experiments, a given feature (e.g. at q_0) will frequently have harmonics (at $2q_0$, $3q_0$, etc.). Interpreting this axis as frequency in a sonification would naturally generate audio overtones which the human auditory system is exceedingly well-equipped to detect: timbre. Timbre is difficult to define, but has been described as "that attribute of auditory sensation in terms of which a listener can judge that two sounds, similarly presented and having the same loudness and pitch, are different." [3]

In this paper, we explore sonification as a tool to provide scientists with an additional method to deeply explore scattering datasets. The abstract nature of the data makes this a challenging, but critical, problem. Moreover the quantity of such data generated is growing hugely with time: newer x-ray instruments are now being built with ever-greater flux, generating data at an ever-increasing speed. It is also worth noting that scattering experiments can also be performed with visible light, electron beams, and even neutron beams. Although we focus very specifically on x-ray scattering data in this paper, we view this as a case study for the general problem of extracting meaning from the highly abstract datasets that are common in the physical sciences. We show that timbral sonification generated directly from the data through additive synthesis [4] can provide a useful instantiation of abstract data.

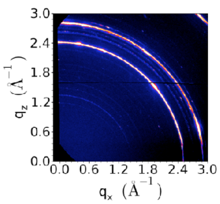

Figure 1: Example x-ray scattering data. The direct beam is incident near the lower-left corner of the image. The false-color image highlights certain features which arise from diffraction of x-rays from the sample's internal structure.

2. SONIFICATION

Over the last few decades, there have been a number of interesting cases of data sonification. Sonifications have been made of seismic data [5], ocean currents [6], and heart rates [7]. Despite these examples, sonification is a largely underutilized technique. Sonification provides a number of unique advantages: the human ear has a wide dynamic range across two variables: frequency and loudness; the human auditory system is finely tuned to detect subtle changes and extract signals from substantial noise; sonifications can be ambient, rather than requiring focused attention; and sonifications can be added to other forms of data exploration, creating more immersive multi-modal interactions.

Much of the existing work in sonification has involved conversion of time-series data. Such conversions are undoubtedly valuable, and are intuitive to understand, but this leaves aside the vast majority of datasets, where some non-temporal variable is of interest. In addition, recent sonifications have mapped the input data onto a tonal scale, or even used sampling or synthesis to reproduce notes from particular instruments. [8] These musical sonifications, like music itself, exploit pattern-seeking features of the human auditory system to create sounds that are crisp, distinct, recognizable, and typically pleasant. Although such realizations can be interesting, even beautiful, the musical nature frequently obscures the underlying patterns in the data. Herein we advocate for the more direct mapping between data space and sound. This necessarily leads to more complex, even cacophonous, sonifications; however such a mapping is relatively unbiased and preserves the majority of the information content. One can crudely identify a tradeoff between aesthetics and information content. Our sonification method uses pitch and loudness only to inform the additive synthesis; the main auditory channel is timbre.

We reformulate the two-dimensional scattering image into a (q, angle) array, where 'angle' is the arc angle with respect to the vertical axis of the image. In so doing, rings of scattering (which have a constant q-distance from the incident x-ray beam) are turned into straight horizontal lines in the $I(q,\text{angle})$ matrix. Doing so also highlights any variation in the ring intensity, which corresponds to spatial orientation of the structures in the sample. The intensity matrix has no time variable; we introduce time by in effect sweeping through the experimental data. In particular, the $I(q,\text{angle})$ matrix is directly converted into an $I(f,t)$ matrix, where f is frequency and t is time. This matrix is simply a spectrogram, or sonogram, which can of course be converted into a sound waveform through additive synthesis. For a sampling rate f_s:

$$A(t) = \sum_{n=1}^{N} I_n(t) \sin\left(\frac{2\pi f_n}{f_s} t\right) \tag{5}$$

here $A(t)$ is the instantaneous amplitude of the output waveform, and the $I(f,t)$ is discredited into $I_n(t)$ by splitting the frequency range into N bins. Thus the scattering data (the $I(q,\text{angle})$ matrix) is mapped directly into the amplitudes of the sine wave components of the sound. This synthesis inherently creates timbre-based (as opposed to tonal) sounds.

We wrote a simple program, using the Python programming language, which directly performs the computation in equation (5), and outputs the resultant waveform into a sound file. We note that this brute-force computation of the waveform is not necessarily the most computationally efficient, or elegant, means of performing additive synthesis (e.g. an appropriate FFT could be used). However we elected to use this method in order to provide flexibility in terms of redefining the mapping between the input data and the output waveform.

The mapping of q into frequency is extremely natural. As already described, both q and f are in some sense the variables along which a Fourier transform is taken. Both exhibit overtones and other natural relationships. The selected mapping is essentially taking the spatial modes (c.f. equations (2) and (3)) and mapping those into frequency modes. Although the one-to-one mapping between the $I(q,\text{angle})$ array and $I(f,t)$ array is information-preserving, and relatively natural, we must make a number of choices about what ranges to specifically map between.

3. PARAMETER OPTIMIZATION

In producing audio files from the two-dimensional data matrices, we must make a number of decisions about both audio encoding, and the range of the mapping (e.g. how to scale between angle and time). A sampling rate of $f_s = 44.1$ kHz (CD audio quality) was selected to provide sufficient quality for the detailed structures in the scattering data. Similarly, a 32-bit intensity encoding was used to allow for the large dynamic range of scattering datasets. As mentioned, there is a natural relationship between q and f. We align $q = 0$ with $f = 0$ so that any harmonics (or other natural progressions) in the scattering data are automatically converted into harmonics in the sound output. Scattering images are typically visualized using a false color map applied through a logarithmic scale, the human auditory system makes this unnecessary for sonification.

Further parameters were optimized by testing a variety of values. For this testing we used scattering data from a polymer solar cell material confined in a nanoscale grating (see Figure 2). Physically, this sample has an oriented morphology; this translates to a scattering ring whose intensity varies along the arc. This, in turn, translates into time variation of the sonification.

The mapping along the frequency axis, which encodes the q-values, is necessarily arbitrary. Although there is a natural reason to align the origins of q and f space, there is no clear correspondence between inverse-distance and inverse-time units. The maximum frequency for human hearing is ~20 kHz. However this choice of frequency maximum was found to generate sounds with too many piercing components. Selecting too low a value for the frequency ceiling resulted in deep and rumbling sounds which essentially washed out all the structure in the scattering data. We found that an upper bound of ~5 kHz in frequency resulted in sonifications that were rich and preserved important data features, without leading to ear fatigue.

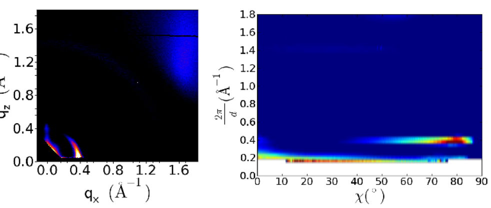

Figure 2: X-ray scattering data in false color on the left. The image on the right remaps the scattering data into an $I(q,\text{angle})$ matrix.

The partitioning of the frequency axis into N bins has a substantial effect on the quality and character of the final sound. An extremely low value (e.g. 10 bins per $Å^{-1}$), not surprisingly, over-smoothed the data and resulted in a loss of data. However, extremely fine partitioning (e.g. 1000 bins per $Å^{-1}$) introduced drastic beating artifacts into the sound. Essentially, by having more frequency resolution than actually warranted by the data's q-resolution, we introduce step-edges in the frequency envelope. The optimized value (50 bins per $Å^{-1}$ for the test dataset) reproduces the spacing in the original data.

The construction of the $I(f,t)$ matrix also requires an arbitrary choice about temporal discretization. Note that this binning width is not the same as the sampling frequency, f_s. Whereas f_s describes the sampling rate used in the additive synthesis (the construction of the output waveform), the temporal binning describes the partitioning of the $I(f,t)$ matrix used to compute the amplitude values for the synthesis. The temporal resolution here is limited by the original dataset. As expected, using low temporal resolution (10 bins per second) smoothed over features in the data, effectively throwing away data. Higher data rates of course cure this defect. However, there is no advantage to increasing the time partitioning beyond that dictated by the initial data. We found that 50 and 1,000 bins per second were found to be essentially identical. We selected 150 bins per second as the optimal value, allowing a healthy safety margin. We improved the sound substantially by interpolating between the data points along the time axis. Doing so avoids sudden changes which introduce sharp popping artifacts into the sound, which hinders comprehension (not to mention damaging speakers).

The length of the sound has a strong effect on the listener's ability to discern structure. Sounds that are too short are difficult to parse. Stretching the sound helps reveal certain details, but inherently makes changes more gradual and difficult to notice. We found that sounds less than 1 second were too fast to be of any use. Sounds on the order of 1-2 seconds could potentially be useful for quick comparisons and identifications, but were still too fast to truly notice signal variations. At 3.5 seconds, sounds, and trends within those sounds, were discernible. Stretching sounds beyond ~10 seconds made it harder to track feature changes.

The above parameter optimization confirms certain limits of the sonification process, but is in some sense idiosyncratic to the datasets chosen. Ideally, all of these variables would be quickly and easily tunable by the user, allowing them to explore datasets in different ways. Looking forward, we envision a software interface that allows the user to select subsets of the scattering data to sonify, and allows the mapping ranges themselves to be easily modified.

4. VARIANTS

In the foregoing, we have attempted to motivate the use of the most direct, perhaps most naïve, mapping between the input data and the final waveform. We also explored a variety of alternative mapping strategies. Imposing additional mapping rules can be a powerful way to highlight certain features of datasets, and this is a valuable way to explore data through sound. We considered the following alternate mapping of intensity to waveform amplitude:

$$A(t) = \sum_{n=1}^{N} \sin\left(\frac{2\pi f_n}{f_s} I_n(t)\right) \tag{6}$$

Here, rather than the intensities modulating the amplitude of the sine waves, they modulate the frequencies of these waves. By using the data matrix to modulate frequency, rather than amplitude, the character of the sound changes substantially. Changes in intensity become very strongly highlighted, as they produce noises that vary in pitch. These chirps or 'boomerang' sounds are distinctive and can be useful for uncovering subtle intensity changes, or small peaks, that might otherwise go unnoticed.

For many samples of interest in x-ray scattering, there is no preferred orientation of the material. Experimentalists typically convert these two-dimensional datasets into one-dimensional curves by averaging overall all possible angles in the image. Sonifying the original two-dimensional data using the approaches described above would result in a sound that does not vary with time. One obvious alternative mapping that we explored is to simply sweep time through the horizontal axis (q), and use the intensity to modulate the amplitude of a single tone at frequency f:

$$A(t) = I(t)\sin\left(\frac{2\pi f}{f_s} t\right) \tag{6}$$

Although simplistic, this mapping can be useful. In particular, the existence of equally-spaced peaks in scattering data yields a metered oscillation in the sound. Moreover, subtle deviations of peak positions could be picked up by the listener, as hearing is able to discriminate small timing differences. As with the two-dimensional data, we can use the intensity data to instead modulate the frequency of the sound:

$$A(t) = \sin\left(\frac{2\pi f}{f_s} I(t)\right) \tag{7}$$

Here again, we discover that by modulating frequency, rather than amplitude, sudden changes in intensity in the data become highlighted by sweeping changes in frequency. Details of peak positions and heights are sacrificed, but extremely weak peaks now become readily apparent. This points again to the need in sonification for user-adjustability.

5. APPLICATIONS

Scientists studying x-ray scattering have already developed a sophisticated toolbox of visualization techniques to explore data, and theoretical models to explain, quantify, and fit their data. It is thus natural to ask whether sonification can bring any new insight to the task of understanding these abstract datasets. We envision a variety of ways in which sonification could elucidate experiments. Consider the data shown in Figure 3, for four different kinds of samples. The false-color images are all quite distinct; and indeed the corresponding sounds are all unique and extremely distinct: the first image has many striations which leads to a number of fairly distinct tones persisting in time. The second image is a 'misaligned' sample; the corresponding sonification is dominated by blips and cracks that sound distinctly like artifacts. The third example is a composite of nanotubes dispersed in an elastic polymer. The scattering image has diffuse intensity throughout, due to the disordered arrangement in the sample; this can be heard as a hazy, wind-like sound permeating the sonification. The final example is a nano-scale grating. Here, the extremely regular and precise structure results in many distinct streaks in the false-color image. These streaks create periodic rhythms in the sonification.

One notable advantage of sonification over careful visual inspection is that the former can be done ambiently. Modern scientific instruments are becoming increasingly automated, to handle the growing scale of scientific discovery. Sonification provides the opportunity for the experimenter to work on other tasks, while listening, in the background, to automated data collection. Any sudden changes in the incoming data, or surprising samples, will immediately be noticed and can be explored in greater detail. Consider for instance the 'misaligned' sample; the sonification is distinct and the experimenter would immediately know that something was wrong with the instrument.

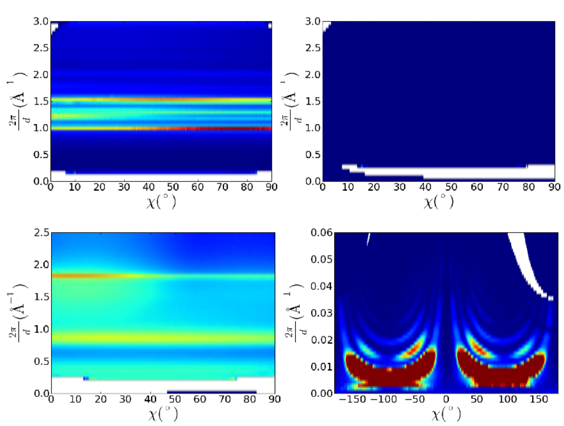

Figure 3: Examples of the variety of data one can obtain from x-ray scattering. From top to bottom the sample are: a semi-crystalline commercial plastic; a 'misaligned' sample (where the beam missed the sample); a composite of carbon nanotubes in a matrix of elastic polymer; and an empty nano-scale grating.

With some effort and training, it is also likely that an experimenter could learn to differentiate between all the unique features in the sound, and could pull out interest trends and features that they had ignored in a visual analysis. It is clear, however, that what is lacking are fast and easy-to-use software tools to enable users to quickly explore different mappings and different datasets.

6. CONCLUSION

We have presented a case study of sonifying x-ray scattering data. Direct mapping of the two-dimensional intensity values of a scattering dataset into the two-dimensional matrix of a sonogram is a natural and information-preserving operation that creates rich sounds. Our work supports the notion that many problems in understanding rather abstract scientific datasets can be ameliorated by adding the auditory modality of sonification. We further emphasize that sonification need not be limited to time-series data: any data matrix is amenable.

Timbral sonification is less obviously aesthetic, than tonal sonification, which generate melody, harmony, or rhythm. However these musical sonifications necessarily sacrifice information content for beauty. Timbral sonification is useful because the entire dataset is represented. Non-musicians can understand the data through the overall color of the sound; audio experts can extract more detailed insight by studying all the features of the sound.

7. ACKNOWLEDGMENT

X-ray scattering experiments were carried out on the X9 beamline, which is managed by the National Synchrotron Light Source and the Center for Functional Nanomaterials, Brookhaven National Laboratory, which are supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC02-98CH10886. We thank Danvers Johnston for providing the test sample for the scattering experiment.

8. REFERENCES

[1] B.E. Warren, *X-Ray Diffraction*, New York, USA: Dover Publications, 1990.

[2] T. Fitch & G. Kramer, G. *Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multi-variate system.* In G. Kramer (ed.), Auditory display: Sonification, audification and auditory interfaces. Proceedings of the First International Conference on Auditory Display (ICAD) 1992, 307-326. Reading, MA: Addison-Wesley, 1994.

[3] J. Neuhoff, "Perception, Cognition and Action in Auditory Displays." In T. Hermann, A. Hunt, J. Neuhoff. (ed.), *The Sonification Handbook.* Berlin: Logos Publishing House, 2011.

[4] C. Roads, *The Computer Music Tutorial,* Cambridge, MA: MIT Press, 1996.

[5] J. Acoust. *Seismometer Sounds.* Soc. Am. Volume 33, Issue 7, pp. 909-916, 1961.

[6] B. Sturm. "Pulse of an Ocean: Sonification of Ocean Buoy Data." *Leonardo*, Vol. 38 Issue 2, pp. 143-149, 2005.

[7] M. Ballora et al., "Heart Rate Sonification: A New Approach to Medical Diagnosis," *Leonardo* Vol. 37, No. 1, pp. 41–46, 2004.

[8] J. Kreidler *Charts Music – Songsmith fed with Stock Charts.* http://www.youtube.com/watch?v=2-BZfFakpzc&feature=player_embedded, 2009.